

O I P E

JUL 02 2002

Verf1: 1 to 3640

## Figure 1

CGGGGGGAGTGGGGAGGGAGGGGGTCGGCCGCCGCAGCCATGGAGGCCAAGCTGGACGCCGTTCCGTGTTCCAGGCCAACGAAGCATCCCAT  
M E A N W T A F L F Q A H E A S H 17 90

CAACAAACAGCAGGCCAGCGCAGAACAGCTGCTGCCCTCTGAGTTCTGTGAGGCCAGCTGATCAAGAAACCGTTGCTTCAATACCA  
H Q Q Q A A Q N S L L P L S S A V E F F D Q K P L P I P 47 180

ATTACTCAGAAACCTCAGGCTGACCAGAACATTAAAGGATGCCATTGGGATTAAGAAAAGAAAAACCCAAAACCTCGTTGTGCACT  
I T Q K P Q A A P E T L K D A I G I K K E K P K T S F V C T 77 270

TACTGCAGTAAAGCATTAGGGACAGCTATCACCTGAGGCCATCAGTCTGCCACACAGGGATCAAGTGGTGTCTGGCAAAGAAA  
X C S K A F R D S Y H L R R H O S C H T G I K L V S R A K K 107 360

ACCCCCACCGTGGTCCCTTATCTCACCATTGCTGGGACAGCAGCCGAACTTGTGTTCAACTATTGCAAGGCATCTGTCA  
T P T T V V P L I S T I A G D S S R T S L V S T I A G I L S 137 450

ACAGTCACTACATCTCCTCGGCACCAACCCAGCAGCAGCGTAGTACCAACAGCACTGCCGTGCCCCAGCTGTCAAGAAACCCAGT  
T V T T S S S G T N P S S S A S T T A M P V P Q S V K K P S 167 540

AAGCCTGTCAAGAAGAACCAACGCTGTGAGATGTGAGGCTTCCGGGATGTGACCACTCAATCGGCACAAGCTCTCCATTG  
K P V K K N H A C E M C G K A F R D V Y H L N R H K L S H 197 630

GACGAAAAGECCTTTGAGTGTCTTATTGTAACTCAGGCCCTCAAGAGGAAGGACCGGATGACTTACCATGAGGTCTCATGAAGGAGGC  
D E K P F E C P I C N O R F K R K D R M T X H V R S H E G G 227 720

ATCACCAAACCCCTATACTTCAGTGTGTTGTGGGAAAGGCTTCTCAAGGCTGACCACCTAAAGCTGTCACTGAAAACATGTGCAACCA  
I T K P Y T C S V C G K G F S R P D H L S C H V K H V S T 257 810

GAAAGACCCCTCAATGCCAACAGTCACGTGCTGCCCTTGCCACCAAAGACAGACTACGGACACACATGGTGCAGCCACAGGAAGGTA  
E R P F K C O T C T A A F A T K D R L R T M V R H E G K V 287 900

TCATGTAACATCTGTGGGAGCTCTGAGTGCAGCATATATCACAGCACTTAAAGACACATGGGAGGCCAAAGTATCACTGTAAAC  
S C N I C G K L L S A A X I T S H L K T H G Q S Q S I N C N 317 990

ACGTGCAAACAAGGCATCACCAAAACGTGCATGAGTGAGGAGACCGCAATCAGAAGCAGCAGCAGCAGCAACAGCAGCAGCAG  
T C K Q G I S K T C M S E X T S N Q K Q Q Q Q Q Q Q Q Q Q Q Q 347 1080

CAACAACAACAACATGTGACAAGCTGCCAGGGAGCAGGTAGAGACACTGAGACTGTGAGAAGCTGTCAAAGCAAGAAAGAAAAGAA  
Q Q Q Q Q H V T S W P G K Q V E T I R L W E E A V K A R K K E 377 1170

GCTGCCAACCTGTGCCAACCTCCACGGCTGTACGACACCAAGTGACTIONTCAACTCCATTCAATAACGTCCTCTGTGCGTCTGGG  
A A N L C Q T S T A A T T P V T L T T P F N I T S S V B S G 407 1260

ACTATGTCAAACCCAGTCACAGTGGCAGCTGCAATGAGCATGAGAAAGCTCAAGTGTGAGCTAACATAACCAAGCCCTTA  
T M S N P V T V A A A M S M R S P V N V S S A V N I T S P L 437 1350

GCCATGACCTCACCTTAAACTCACCCACCCAGTCACCTCCCCACCCCTGTGACCGCCCAAGTGAATATAGCACACCCGTCAACCATC  
A M T S P L T L T T P V N L P T P V T A P V N I A H P V T I 467 1440

ACATCTCCAATGAACCTGCCACTCTATGACATTAGCTGCCCTCTCAATATAGCAATGAGGCTGTAGAAAGTATGCCCTCTGCC  
T S P M N L P T P M T L A A P L N I A M R P V E S M P F L P 497 1530

CAAGCTTGCCATCGTCACGCCCTGGTAAACAGTATTATAAGTCAAAATTGGGTTAAAGTAAATATTACAGCAACTAACCTTAGTT  
Q A L P T S P P W 506 1620

GATTAAGCAAAAGCAGACTATGAAATTGGGGTTTATATGTTAGTTAAAGAGTGTAGTGTGTCCTAATTGCTGGGGTGTTC  
AAAGTAGGGTATATGTTAACTTATCACTGGACCACTTTAGTTACTCAGAAACCCCTTAGCTGACACCATCTGTTAACAGGATAGTA  
GCTGGCAAGGAAATGCCGAAGTAAACCAATCAAAACCCATTCTGTTAACAAATAAAAGCATTATTGTTTTTATTATTTTTAAAT  
ACAACAGAAATCTTTATTGTAAACACTAGCAGAGTCTCCCTCTGTACAGGTGGACGGTTAACCTGGAGCTCAAGCCACAGACT  
GAGAGCTAGTGTAGCATGTGTTGCTGTAGTGAACAGAGGCAATTGCTATAAAATGCAATTTCAGAGATAATGCAATT  
TACCTTTGGGAATATGTTAACTTCAGGCAGATTCCCTATGGGAAAGGTGATACCGACTCTGTATATGCAAAAGCATATGATAATTATCAT  
TCTAACTTCAACATATAATAGGGATGTGACCTGATATTGGAGATGTAAATTAGGCTGACATATTAACTCTGTGATGATATAGCATT  
GTAGTTGACTTTTAAACAAAACAAAACAAAACAAAAGGAATMCCGAGAGCTGTCACTGCCATTGAAATGCTATCCCTGGGATAGG  
GGTGGCTTCAGAACCCAGGAAGTGGCAAGGGCACAGACTCTGCTGGAGGCCGTGAGCCGGGGTTCCATAGGAGACTGACAGGAGACAT  
TTGGCTTGGAGAACAAAAGAGNAGGCTACCCCACTTCAAGATGAGCAGACAGGCAATTGCTGAGCTGAGTTGTTACTGCTGGGATT  
AGTGGTGGAGCAGATGGTACTCTGAGCACTGCTCACAGGCAAGGCTGAGCTGAGCTGAGTTGTTACTGCTGGGATT  
GACCTTTGTGCAAGCGTGTATCACGCCCTACCTCACCACTGCGACGGACACAGAGCTGAGCACCCTACCGAACCATCTACCGACTGC  
CTATCGCCGTAGGCCCTGGGGTACTCCCGCAAGGCCCTCGTGTGCTGCCCTGGTTGAGGAAGGAGCACAGCTGGGCTCCCTGGGCTMG  
TGGAGCAGCAATATGCCACGCCCTATGTGGGAATGGAGGGAGTICATCCGCCAGGACACTGCCGTGCTGGATGGAGATA  
CTTGGCAGACAGATGTTGATGAATGCACTACAGGAGAGGCCAGTGTGCTCCAGCGCTGTGCAATACTGTGGGAAGTACTGTGCG  
GAATGGGAGGGACAAGGCCACTCTGCAAGATGGGACCCCTGCTGTCTAAGGAGGGCCCTCCCTTCCCCAAAACCCACAGCAGGAG  
TGGACAGCATGGGAGAGAGGAGGGTGTACAGGCTGAGCTGGGAGGCTGAGCTGAGCTGAGCTGGGACTCTGAGCTGGGACTCTG  
ACAGCCTGCCCTCGGGTCCACAGAGCATGGGCTACAAAGATCTGGCAGCTGCTGGTGTGCTTCTGAGGAGAACATCTGGGCTCTG  
AGTCCCACAGAACCCACTCCACCCACAGAGAGCCTAGGGGACCCATGGGGTGGACACCAAGGGCTGGGTGAGTGAACCTCTCTGGGAT  
GGGAGATTGCAAGTTTACCCCTCTCTGTGCTCCCTAGGGCTCCCTGAGGAGTAAACCTCTCACCCACAGGCTGGA  
TAGAGCACTACCCAGATCCCTGTAGGCCAGGTTCAAGGGAGCCGTGCTGGTGTGCTATGAGCAGAAGGCCCTGCCCTATTGCTCCCT  
CTTGTGAGGTTCTCTAGGACTGGGTATGGGAGTGGGGTCTTGTGACTCTTCAGTGGGCTCCCTGTCAAGTGGTAAGTGGGAT  
TGTCTCCATCTTGTCAATAAAAGCTGAGACTTGAAGAAAAAA

COPY OF PAPERS  
ORIGINALLY FILED

**Figure 2**

## Human DB1 DNA and Protein Sequences

10            20            30            40            50            60  
 AGCGGGGGGAGTGGGGAGGAGGGGGGTCGGCCGCCGCAGCCATGGAGGCCAACTGGACCG  
 M E A N W T>  
 70            80            90            100          110          120  
 CGTTCTGTTCCAGGCCATGAAGCTTCCCACCAACAGCAGGCAGCACAGAACAGCT  
 A F L F Q A H E A S H H Q Q Q A A Q N S>  
 130          140          150          160          170          180  
 TGCTGCCCTCCTGAGCTCTGCCGTGGAGCCCCCTGATCAGAAACCATTGCTTCCAATAC  
 L L P L L S S A V E P P D Q K P L L P I>  
 190          200          210          220          230          240  
 CAATAACTCAGAAACCTCAGGGTGCACCAAGAAACATTAAAGGATGCCATTGGGATTAAAA  
 P I T Q K P Q G A P E T L K D A I G I K>  
 250          260          270          280          290          300  
 AAGAAAAACCCAAAACCTCATTGTGTGCACCTACTGCAGTAAAGCTTCAGGGACAGCT  
 K E K P K T S F V C T Y C S K A F R D S>  
 310          320          330          340          350          360  
 ATCACCTGAGGCGCCACGAATCCTGCCACACAGGGATCAAGTTGGTGTCCCAGGCCAAAGA  
 Y H L R R H E S C H T G I K L V S R P K>  
 370          380          390          400          410          420  
 AAACCCCCACCACGGTGGTCCCTTATCTCTACCATCGCTGGGACAGCAGCCGAACCTT  
 K T P T T V V P L I S T I A G D S S R T>  
 430          440          450          460          470          480  
 CGTTGGTCTCGACCATTGCAGGCATCTGTCAACAGTCACTACATCTTCCCTGGGCACCA  
 S L V S T I A G I L S T V T T S S S G T>  
 490          500          510          520          530          540  
 ACCCCAGTAGCAGTGCCAGCACACAGCTATGCCAGTGACCCAGTCTGTCAAGAAACCCA  
 N P S S S A S T T A M P V T Q S V K K P>  
 550          560          570          580          590          600  
 GTAAGCCTGTCAAGAAGAACCATGCTTGTGAGATGTGTGGGAAGGCCTCCAGATGTGT  
 S K P V K K N H A C E M C G K A F R D V>  
 610          620          630          640          650          660  
 ACCATCTCAATCGACACAAGCTCTCCATTCAAGATGAGAAACCCCTTGAGTGTCTATTT  
 Y H L N R H K L S H S D E K P F E C P I>  
 670          680          690          700          710          720  
 GTAATCAGCGCTTCAAGAGGAAGGACCGGATGACTTACCATGTGAGGTCTCATGAAGGAG  
 C N Q R F K R K D R M T Y H V R S H E G>  
 730          740          750          760          770          780  
 GCATCACCAAACCCATACTTGCACTGTTGTGGAAAGGCCTCTCAAGGCCTGACCACT  
 G I T K P Y T C S V C G K G F S R P D H>

**Figure 2 (con't)**

790            800            810            820            830            840  
 TAAGCTGTCATGTAAAACATGTCCATTCAACAGAAAGACCCTTCAAATGCCAACGTGCA  
 L S C H V K H V H S T E R P F K C Q T C>  
 850            860            870            880            890            900  
 CTGCTGCCTTGCCACCAAAAGACAGACTGCGGACACACATGGTGCGCCATGAAGGCAAGG  
 T A A F A T K D R L R T H M V R H E G K>  
 910            920            930            940            950            960  
 TATCATGTAACATCTGTGGGAAGCTCCTGAGTGCAGCATACTCACCAAGGCCACTTAAAGA  
 V S C N I C G K L L S A A Y I T S H L K>  
 970            980            990            1000          1010          1020  
 CTCATGGGCAGAGCCAAGTATCAAACGTAAATACATGTAAACAAGGCATCAGTAAAACAT  
 T H G Q S Q S I N C N T C K Q G I S K T>  
 1030          1040          1050          1060          1070          1080  
 GCATGAGTGAAGAGACCAGTAACCAAAAGCAGCAGCAGCAGCAGCAGCAGCAACAAACAC  
 C M S E E T S N Q K Q Q Q Q Q Q Q Q Q Q Q>  
 1090          1100          1110          1120          1130          1140  
 AACAAACAACATGTGACAAGCTGGCCAGGGAAAGCAAGTAGAAACACTCAGACTGTGGGAAG  
 Q Q Q H V T S W P G K Q V E T L R L W E>  
 1150          1160          1170          1180          1190          1200  
 AAGCTGTTAAAGCAAGGAAGAAAGAAGCTGCTAACCTGTGCCAACCTCCACGGCTGCTA  
 E A V K A R K K E A A N L C Q T S T A A>  
 1210          1220          1230          1240          1250          1260  
 CGACACCTGTGACTCTCACTACTCCATTCACTATAAACATCCTCTGTGTCGTCTGAGACTA  
 T T P V T L T T P F S I T S S V S S E T>  
 1270          1280          1290          1300          1310          1320  
 TGTCAAACCCAGTCACAGTGGCAGCTGCAATGAGCATGAGAAGTCCAGTAAATGTTCAA  
 M S N P V T V A A A M S M R S P V N V S>  
 1330          1340          1350          1360          1370          1380  
 GTGCAGTTAACATAACCAGCCCAATGAACATAGGGCATCCTGTAACTATAACCAGTCCAT  
 S A V N I T S P M N I G H P V T I T S P>  
 1390          1400          1410          1420          1430          1440  
 TATCCATGACCTCTCCTTAAACACTCACTACCCCCAGTCAACCTCCCCACCCCCGTCACTG  
 L S M T S P L T L T P V N L P T P V T>  
 1450          1460          1470          1480          1490          1500  
 CCCCAGTGAATATAGCACACCCTGTCACCATCACATCTCAATGAATCTACCCACACCTA  
 A P V N I A H P V T I T S P M N L P T P>  
 1510          1520          1530          1540          1550          1560  
 TGACATTAGCCGCCCTCTCAATATAGCAATGAGACCTGTAGAGAGCATGCCTTCTTGC  
 M T L A A P L N I A M R P V E S M P F L>

**Figure 2 (con't)**

1570        1580        1590        1600        1610        1620  
CCCAAGCTTGCGCTACATCACCGCCTGGTAAACAGTATTATAAAAATCAAAATATGGGTA  
P Q A L P T S P P W \* >

1630        1640        1650        1660        1670        1680  
AAAGTAAAATATTACCGAACTTAACCTTTAGTTGATTAAAGCAAAAAGTAAACCATGA

1690        1700        1710        1720        1730        1740  
AATTGGGAGATTTATTACATTAGTTAATAAGAGTGTGGTAGCATTCTCCAATTGCG

1750        1760        1770        1780        1790        1800  
CTGGGATTATTCAAAGTAGGGTGTGTATGTAACCTACTGGGACCACTTAGTTAATC

1810        1820        1830        1840        1850        1860  
AGAAATTCCCTTTAGCTGACAACATTGCTAACAGGGATAGTAGTTGGCAAGATGAAATG

1870        1880        1890        1900        1910        1920  
CCAGAATTAAAACCAATCATAGTAGAACCCACTTCAAAATAAAAAACAGCATTACTAT

1930        1940        1950        1960        1970        1980  
TTCTAATCCCCAGGAATCACTTATTGTAACACTAGCAGAACTCTCTCCCTATACAAG

1990        2000        2010        2020        2030        2040  
GTGGATGGCTGATTTAACCTGAAATTAAATCCACAGATTGAGAGCTAGTGTAGAATT

2050        2060        2070        2080        2090        2100  
GTCTGTGTTATTGTTTATGAGTAAATACATGCATTGTCATAATAAAATGCATTTCAAG

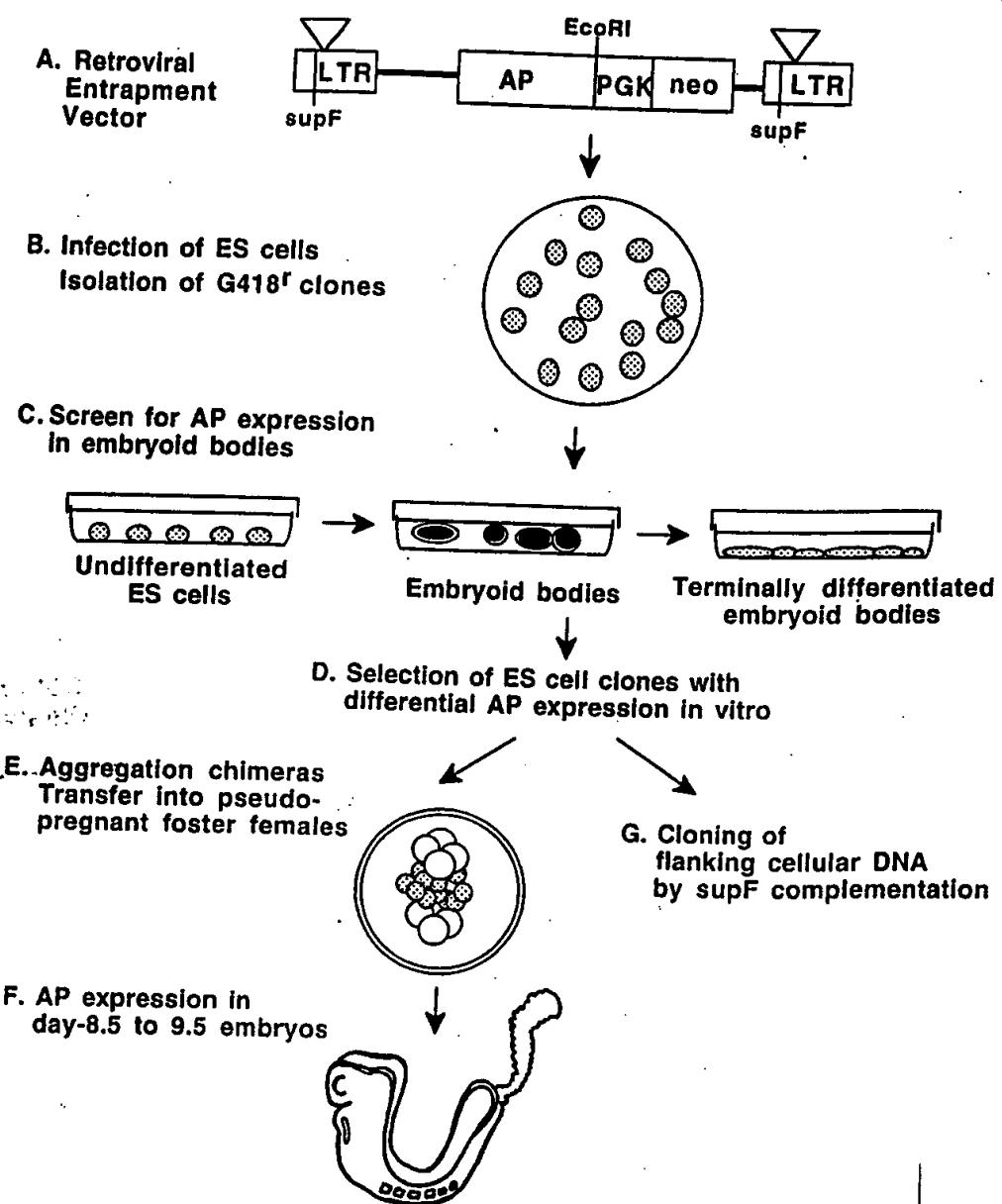
2110        2120        2130        2140        2150        2160  
AGAATATGCATTTACCTTGGAAATATGTTAATTTCAGGCAGCATTCCCTATGGGAAAG

2170        2180        2190        2200        2210        2220  
GTGATACAGCTCTGATATGCAAAGCATATGATAATTATCATTCTAACCTAACGTATA

2230        2240        2250        2260        2270        2280  
ATAGGGATTGTGACCTGATATTGGAGATGTAATATTGCTCAGCATATTAATCCCGATG

2290        2300  
GAATATAGCATTGTAGTTGACTTTT

**Figure 3**



Figur 4

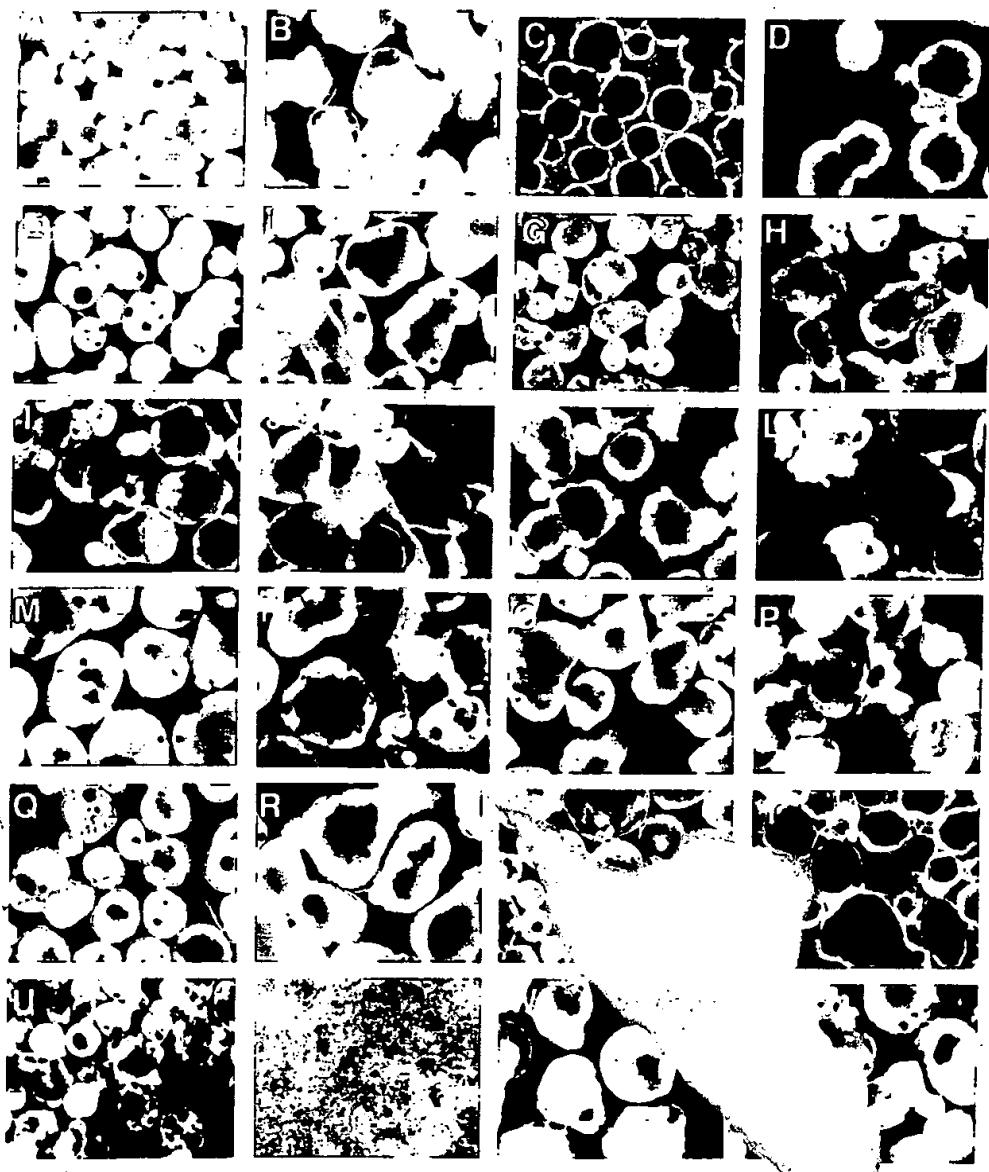


Figure 5

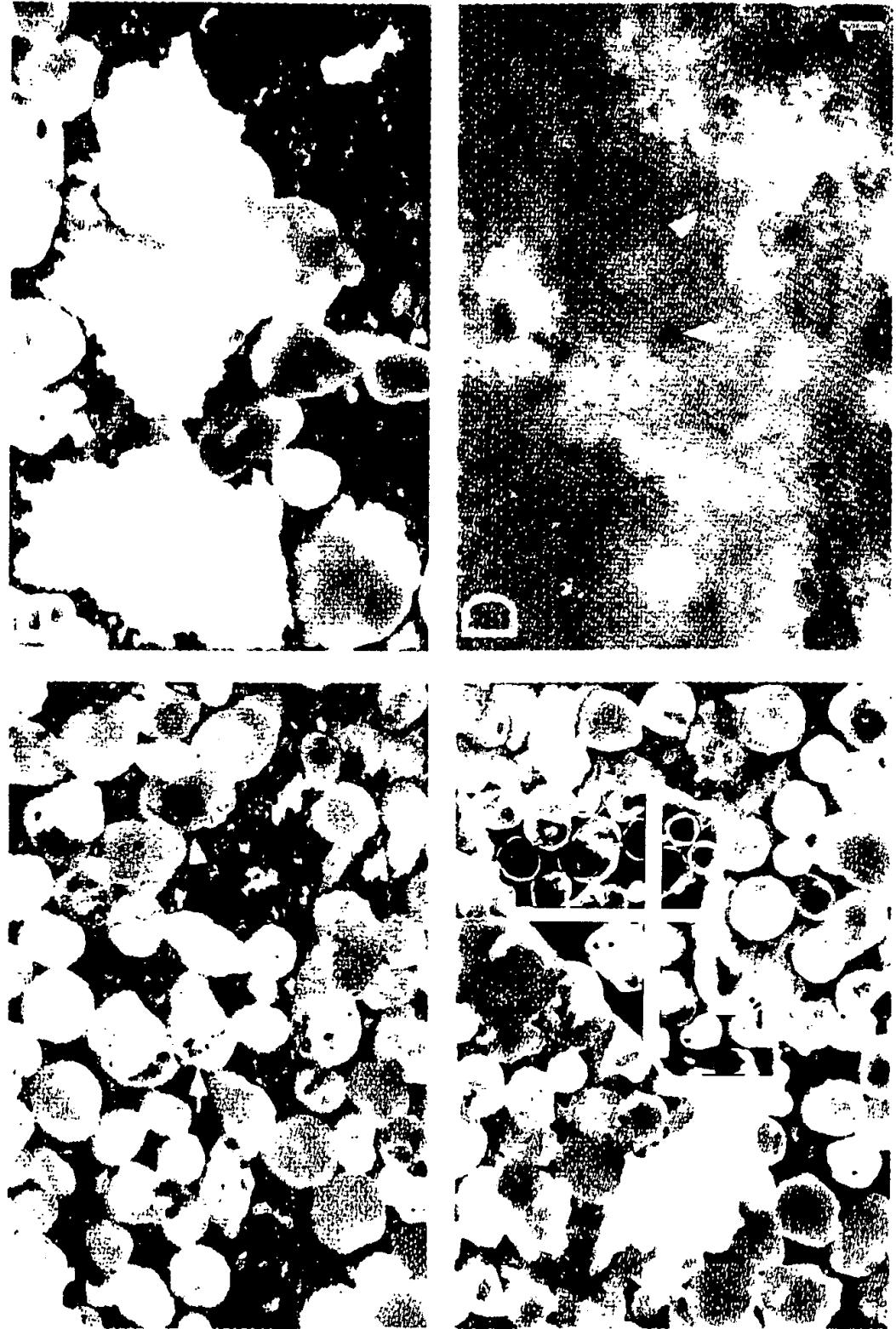
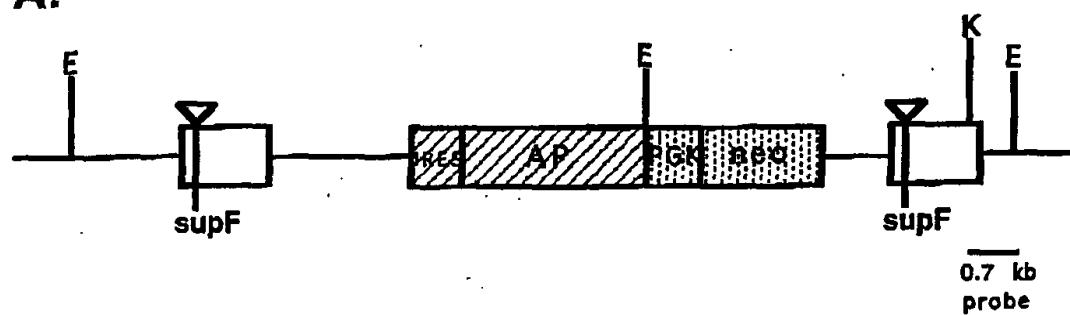
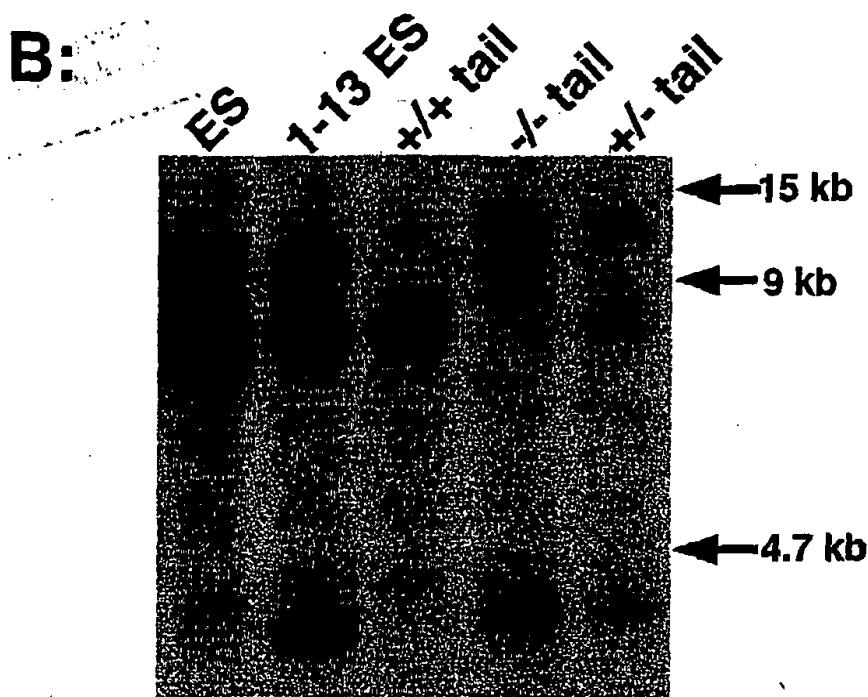


Figure 6

A:



B:



Alignment of Vesf1/mPuri:

Vesf1	168	KVVKKQHACEMOKKAFKUVTILAKKLUHSDEKEFICICHQKFRKQRTTHAVSHBEGITMKPCSCGKFSRPHUSCHRVTHPHTSPRKCQ	264
mPuri	275	KLIRKKAACMOKKAFKUVTILAKKLUHSDEKEFICICHQKFRKQRTTHAVSHBEGITMKPCSCGKFSRPHUSCHRVTHPHTSPRKCQ	372
Vesf1	265	TCTAAFKTOKRSHMUTHECKSVENIGQULLSMTATSHLAKHGSOSLNCTCKQISKTCMSEETSHPKOQQQQQQQQQQQQQHVTSPKGQ	360
mPuri	373	KCEAAFKTOKRSHMUTHECKSVENIGQULLSMTATSHLAKHGSOSLNCTCKQISKTCMSEETSHPKOQQQQQQQQQQQHVTSPKGQ	469

Figure 7

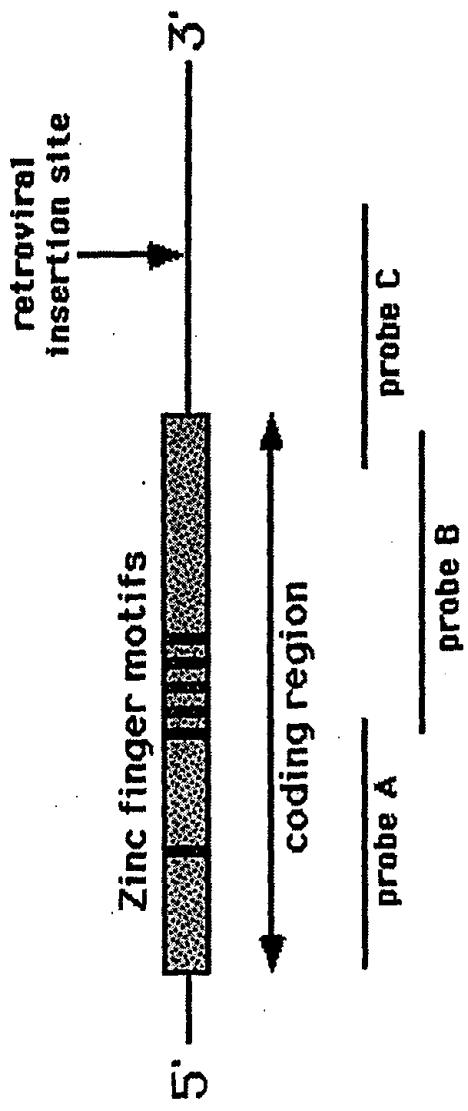


Figure 8

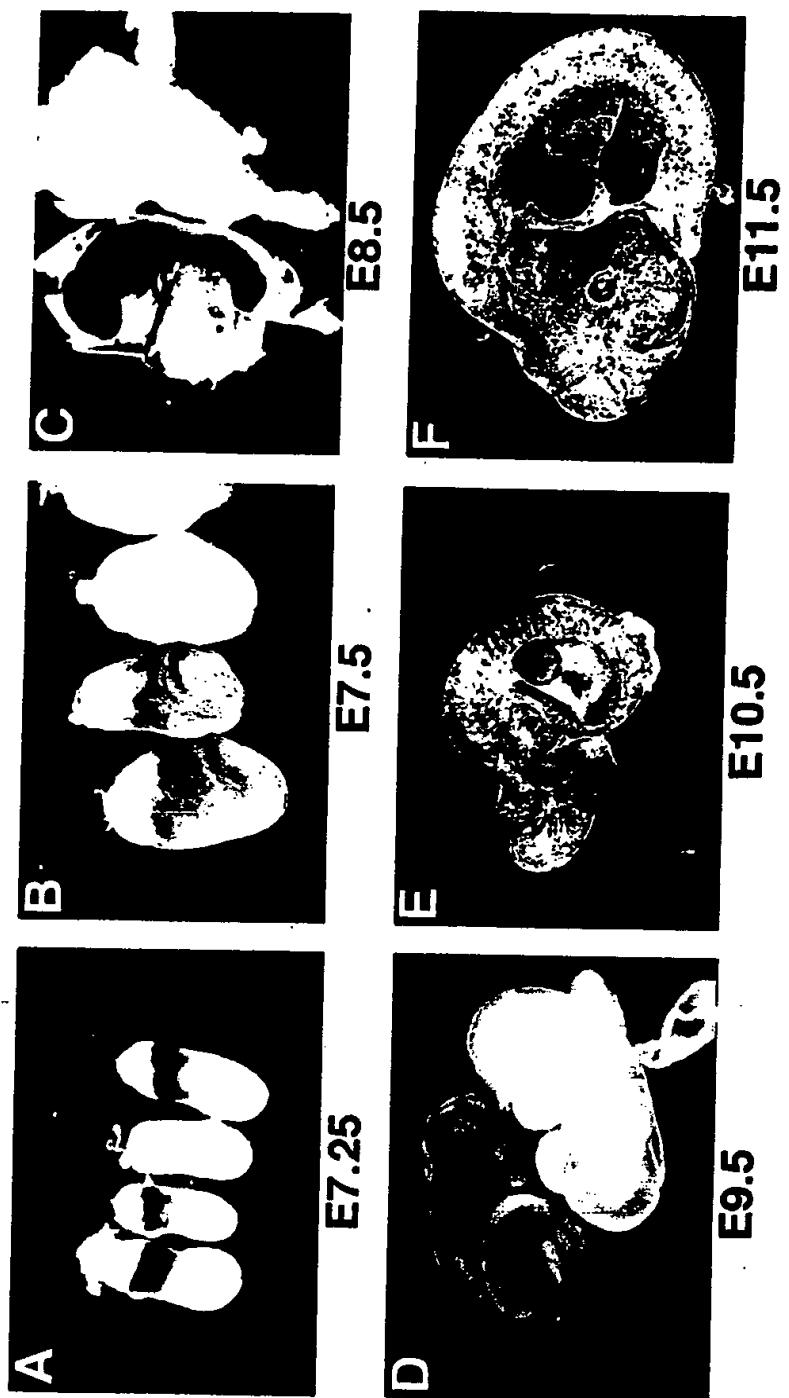
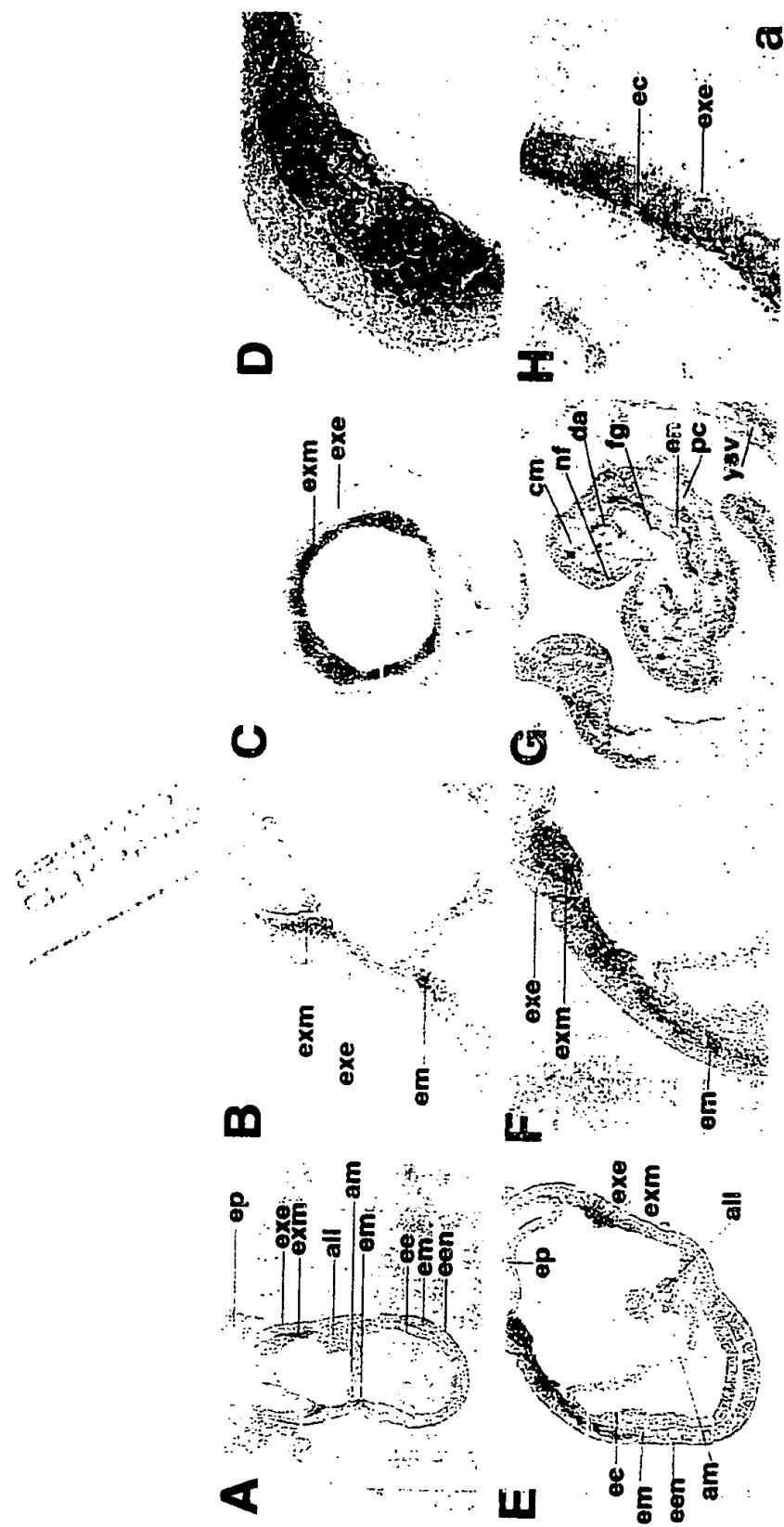
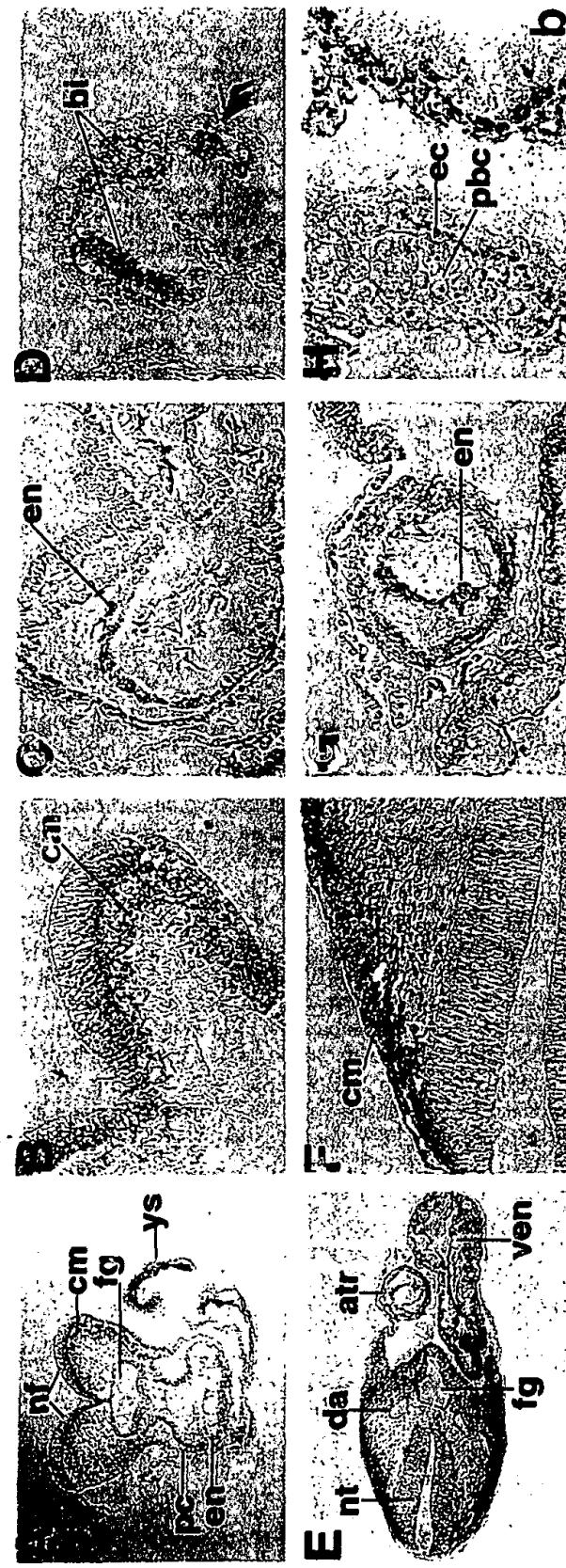


Figure 9

Figure 10



**Figure 11**



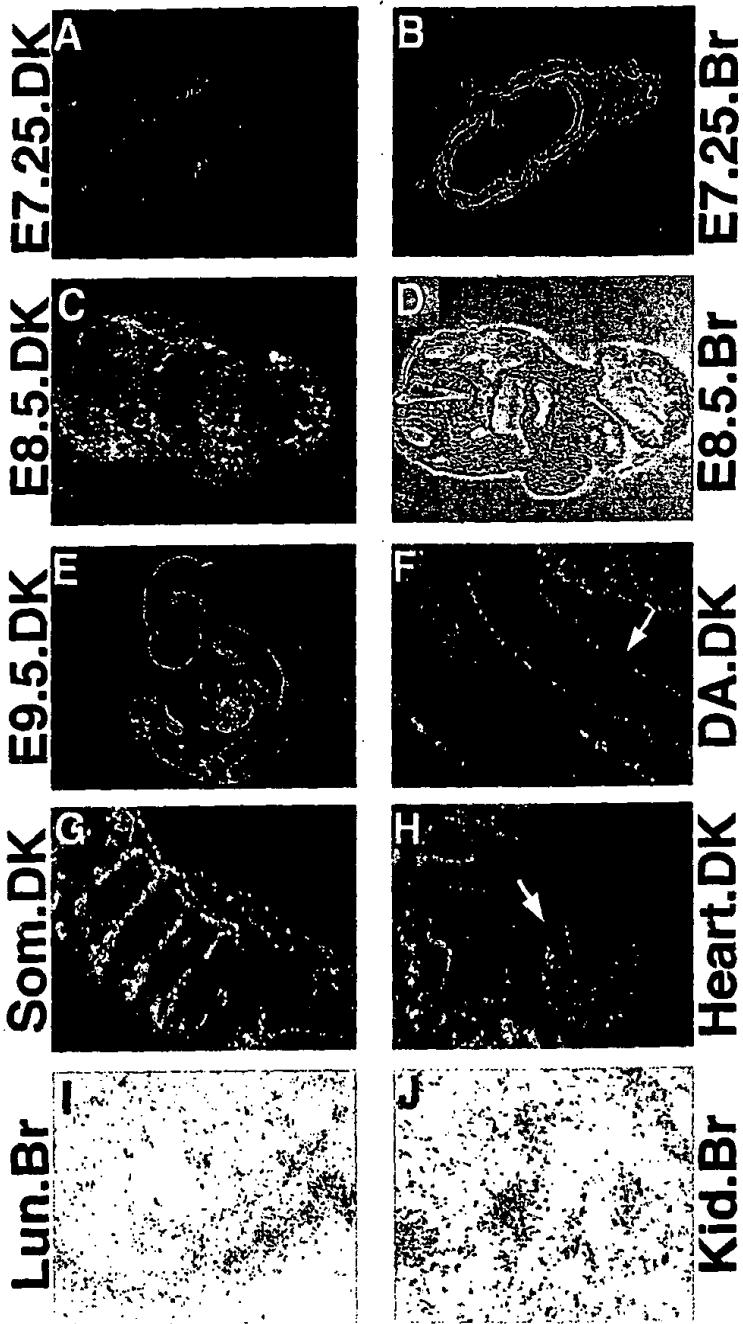
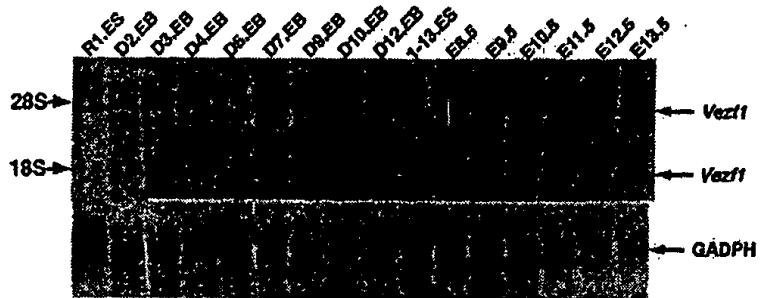


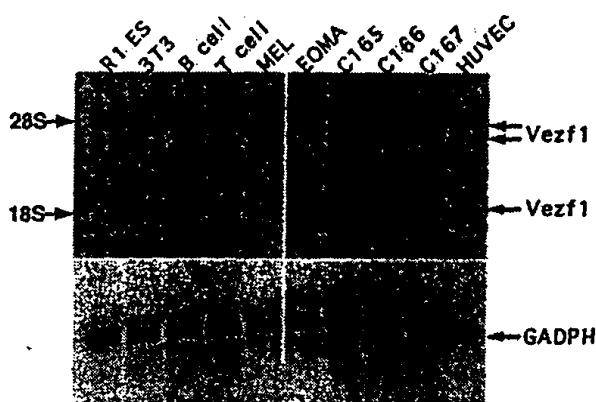
Figure 12

**Figure 13**

**A.**



**B.**



**C.**

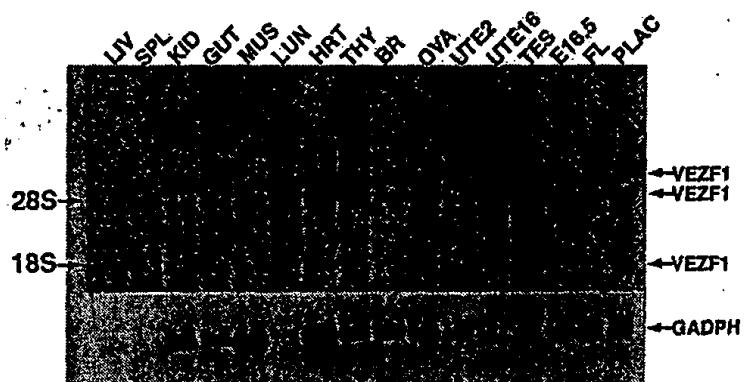
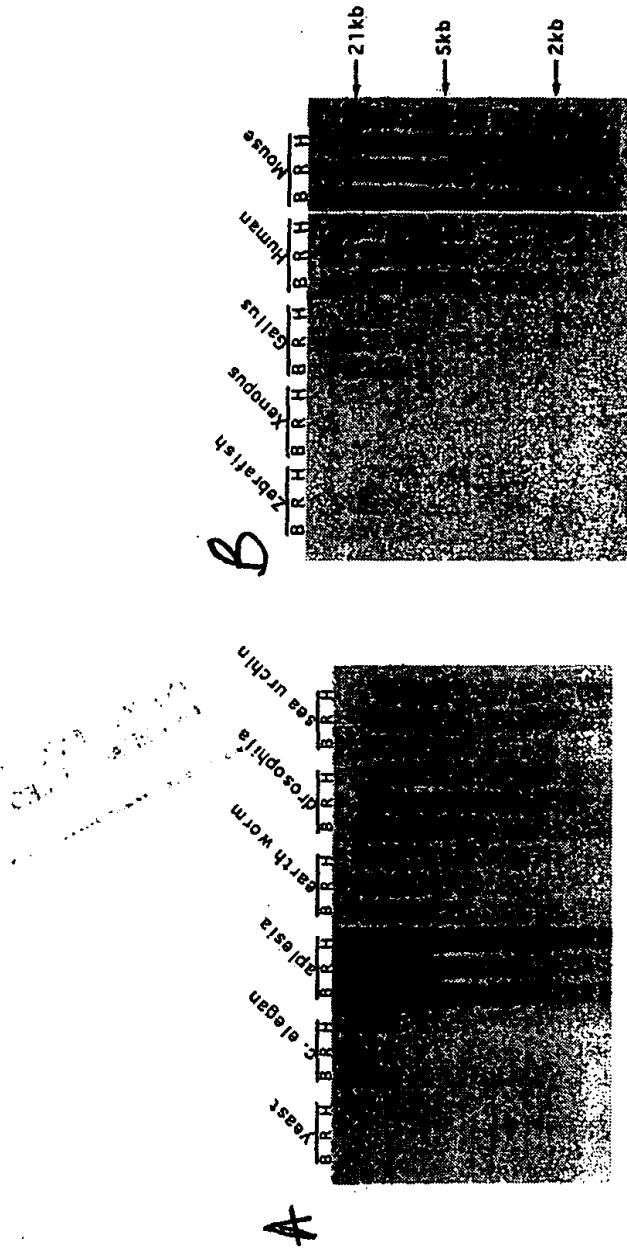
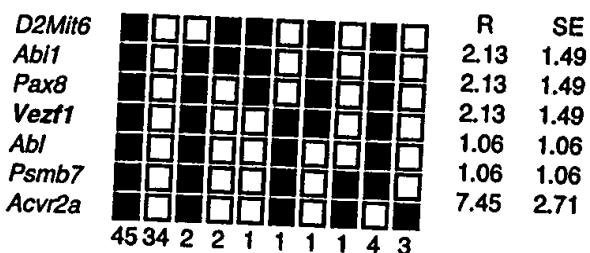


Figure 14



Figur 15

A: Jackson BSS Chromosome 2



B: Jackson BSS Chromosome 2

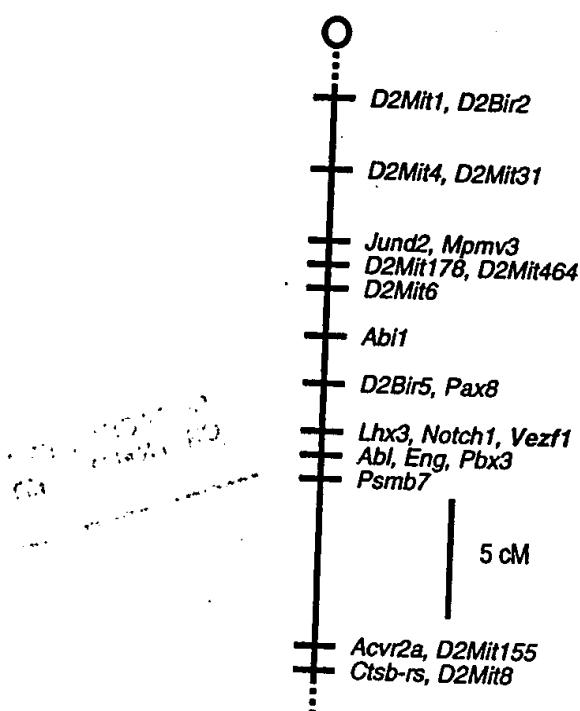
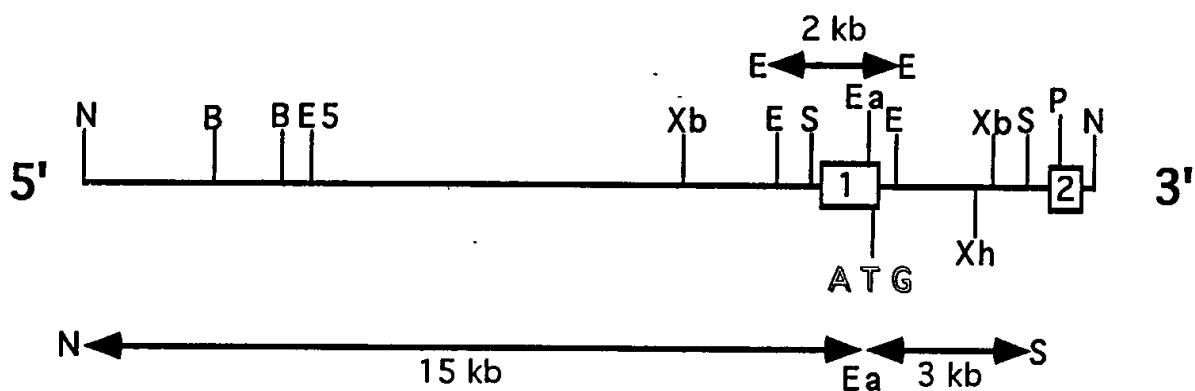


Figure 16

Restriction Enzyme Map of a 20 kb Genomic DNA of the Vezf1 Gene



BamHI (B), EcoRI (E), EcoRV (E5), Eagl (Ea), Notl (N), Patl (P), SacI (S), XbaI (Xb), and Xhol (Xh).

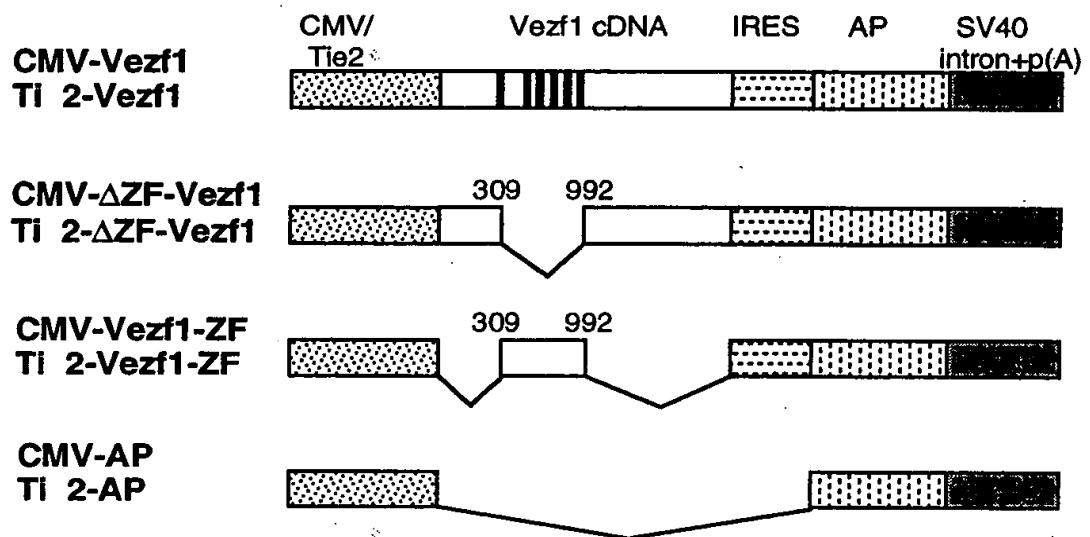
— Intronic sequence;

**[1]** Exon 1

**[2]** Exon 2

**Figure 17**

## **Vezf1 EXPRESSION VECTORS**



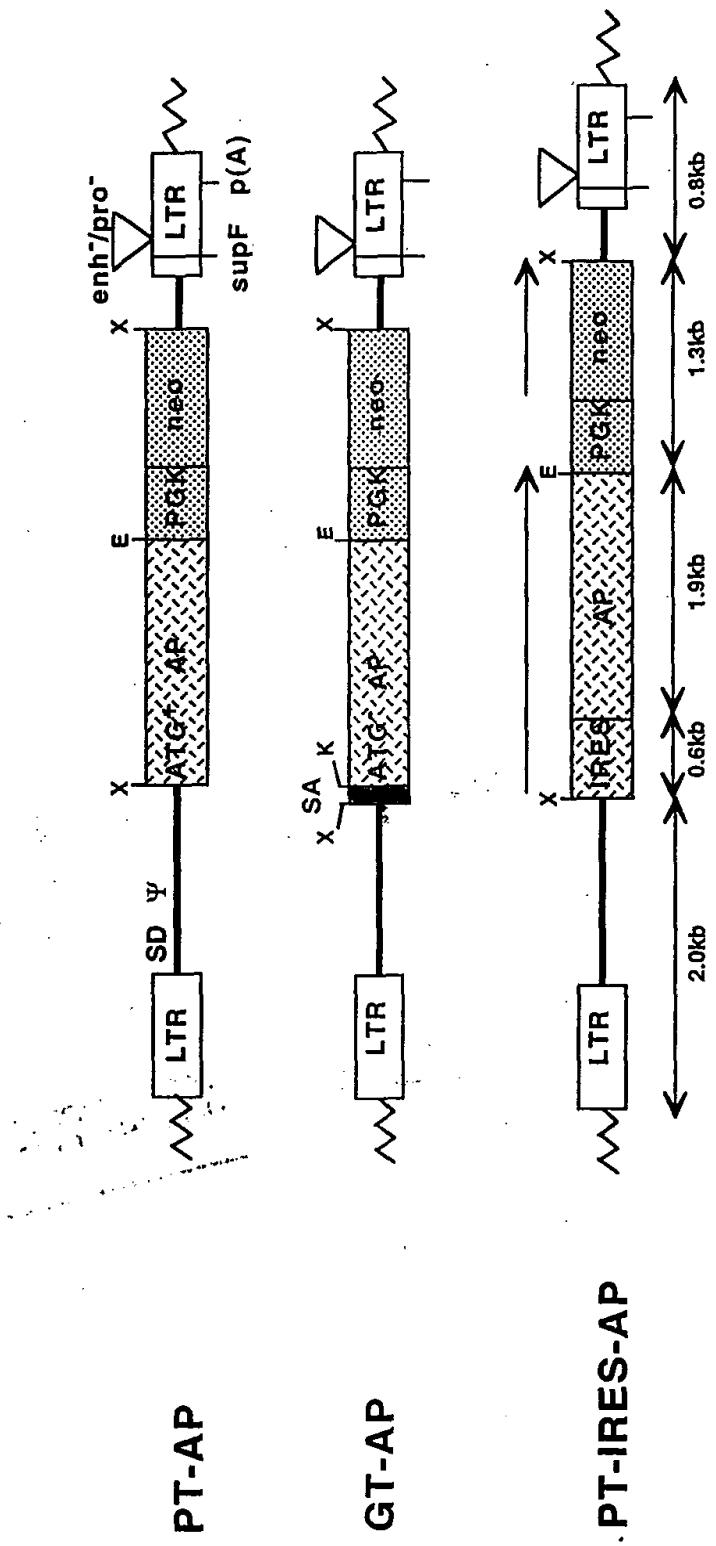


Figure 18